JOURNAL OF MATHEMATICAL ANALYSIS AND APPLICATIONS 28, 303-312 (1969)

Observation and Prediction for the Heat Equation

V. J. MIZEL AND T. I. SEIDMAN

Department of Mathematics, Carnegie-Mellon University, Pittsburgh, Pennsylvania 15213

Submitted by Richard J. Duffin

1. Consider an insulated uniform rod with unknown temperature distribution. With proper normalization the temperature $u = u(t, x)$ satisfies

$$u_t = u_{xx} \qquad (t > 0, 0 < x < 1),$$
$$u_x(t, 0) = u_x(t, 1) = 0. \tag{1}$$

Suppose it possible to observe $f(t) = u(t, 0)$ for $0 < t < T$. We ask: Is it possible, given f, to determine $w(x) = u(T, x)$ for $0 < x < 1$? Assuming the answer to this first question is "yes", is this a "well-posed" problem?

This problem forms a natural sequel to the problem considered in [2]. In that paper it was shown that if $w_0(x) = u(0, x)$ is known and a number $T_1 > 0$ is given then it is possible—by controlling $f(t) = u(t, 0)$ (or $g(t) = u(t, 1)$) over an arbitrarily small time interval $(0, \epsilon)$—to have the temperature distribution $w_1(x) = u(T_1, x)$ approximate any desired $h \in L^2[0, 1]$ arbitrarily closely in norm.

The results of the present study will imply that even without prior knowledge of the initial temperature distribution w_0 it is possible, merely by observing $f(t) = u(t, 0)$ over an interval $[0, T]$ during which (1) holds, to reduce the situation to that considered in [2]. Then, once $u(T, x)$ is known, f and g can be controlled over $(T, T + \epsilon)$ so that $w_2(x) = u(T + T_1, x)$ will again approximate any $h \in L^2[0, 1]$ arbitrarily closely.

The present results are incomplete in two respects. First, in contrast to [2] we have not been able in the present work to extend our results to general spatial regions in R^n, $n > 1$ (however, see the comments concerning cylindrical regions). Second, our arguments indicate that there may be a "minimum observation time" $T_{\min} > 0$ such that the results obtained here are valid for times $T > T_{\min}$ and fail for times $T < T_{\min}$, yet we have not succeeded in determining whether such a minimum observation time actually exists.

We may formulate the present problem more precisely as follows. If (1) holds for $0 < t < T$ then the boundary datum f must satisfy certain consistency conditions, i.e., must lie in a certain manifold $\mathscr{M}$. We are thus asking first whether the operator $A : f \mapsto w$ is well-defined from $\mathscr{M}$ to $L^2(0, 1)$, and second whether it is continuous, topologizing $\mathscr{M}$ by the L^2 norm on $(0, T)$:

$$\|f\|^2 = \int_0^T |f(t)|^2 \, dt.$$

That A *is* well-defined—for *any* $T > 0$—follows from the fact that the only solution to (1) which satisfies in addition the condition

$$u(t, 0) = 0 \qquad 0 < t < T,$$

is $u \equiv 0$. This is a consequence of the unique extension property for the heat equation. (An analogous argument shows that A is also well-defined for a region in R^n, $n > 1$, provided that we observe data on an open portion of the boundary, see [2].) In order to investigate the continuity of A we proceed as follows.

As is well known, we may write the solution of (1) as

$$u(t, x) = \sum_{n=0}^{\infty} c_n e^{-n^2\pi^2 t} \cos n\pi x \tag{2}$$

with appropriate coefficients $\{c_n : n = 0, 1,...\}$ such that

$$\sum_n |c_n|^2 e^{-n^2\pi^2 t} < \infty$$

for $t > 0$. With the substitution $s = e^{-\pi^2 t}$, we have

$$f(t) = \sum_n c_n e^{-n^2\pi^2 t} = \sum_n c_n s^{n^2} = \varphi(s)$$

for $0 < t < T$ or $\alpha = e^{-\pi^2 T} < s < 1$. Then

$$\begin{aligned} Af = w(x) &= \sum_n c_n \alpha^{n^2} \cos n\pi x \\ &= \sum_n l_n(\varphi) \, \alpha^{n^2} \cos n\pi x \end{aligned}$$

where each l_N $(N = 0, 1,...)$ is the linear functional—assuming it is well-defined—such that

$$l_N\Big(\sum_n c_n s^{n^2}\Big) = c_N .$$

Letting

$$[\Lambda] = sp\{s^{n^2} : n = 0, 1, \ldots\},$$

we have, supposing each l_N to be continuous on $M = [\Lambda] \subseteq L^2(\alpha, 1)$, the following estimate (recall that $\mathscr{M} \subset L^2(0, T)$, not $L^2(\alpha, 1)$):

$$\| A \|^2 \leqslant \pi^2 \left(\| l_0 \|^2 + \tfrac{1}{2} \sum_1^\infty \alpha^{2n^2} \| l_n \|^2\right) \leqslant \pi^2 \sum_n \alpha^{2n^2} \| l_n \|^2. \tag{3}$$

What is needed to deduce continuity for A, therefore, is a suitable estimate for the $\| l_n \|$. The Hahn-Banach Theorem assures us that if

$$s^{N^2} \notin [\overline{\Lambda_N}] = \bar{sp}\{s^{n^2} : n = 0, 1, \ldots, n \neq N\} \tag{4}$$

then l_N exists as a continuous linear functional on M with

$$\| l_N \|^{-1} = | s^{N^2} - [\Lambda_N] | \tag{5}$$

where $| \cdot |$ is the L^2 norm on $(\alpha, 1)$, so that

$$\| A \|^2 \leqslant \pi^2 \sum_n \alpha^{2n^2} | s^{N^2} - [\Lambda_N] |^{-2}. \tag{3'}$$

The problem is thus reduced to showing that (4) holds, so that each l_N is well-defined, and obtaining a lower bound for (5).

If we had $\alpha = 0$ then Müntz' Theorem (for which see, e.g., [3]) ensures condition (4) since $\sum n^{-2} < \infty$. For $\alpha \in (0, 1)$, a result due to Clarkson and Erdös [1] shows that $[\Lambda_N]$ is not dense in $L^2(\alpha, 1)$ but does not yield (4), much less a lower bound for (5). We shall obtain a sharpened form of the Clarkson-Erdös result, for a certain class of sequences including $\{n^2 : n = 0, 1, \ldots\}$, showing that (4) holds if α is sufficiently close to zero (i.e., for large enough T) and that then $\| l_N \| = \mathscr{O}(N \log N)$ so that the right hand side of (3) converges and A is continuous.

2. Let $\Lambda = (\lambda_0, \lambda_1, \ldots)$ be a sequence of non-negative reals with $0 \leqslant \lambda_0 < \lambda_1 < \cdots$; let $\Lambda_N = \Lambda \backslash \{\lambda_N\}$. Let $[\Lambda] = sp\{s^{\lambda_n} : \lambda_n \in \Lambda, n = 0, 1, \ldots\}$. We assume that

$$\sum_\Lambda \lambda_n^{-1} < \infty \tag{6}$$

and set

$$\tau_N = \sum_{n > N} \lambda_n^{-1}.$$

It is convenient to introduce

$$r_n{}^N = \begin{cases} \lambda_n/\lambda_N & n < N \\ \lambda_N/\lambda_n & n > N \end{cases}, \qquad r_N = \max_n\{r_n{}^N : n \neq N\} < 1$$

and

$$\varphi(r) = -\log[(1-r)/(1+(1+\lambda_1^{-1})\,r)].$$

If we set, for n, $N = 0, 1,\ldots$ with $n \neq N$,

$$\varphi_n{}^N = \begin{cases} -\log[(1-r_n{}^N)/(1+r_n{}^N+\lambda_N^{-1})] & n < N, \\ -\log[(1-r_n{}^N)/(1+r_n{}^N+\lambda_n^{-1})] & n > N, \end{cases}$$

then

$$0 < \varphi_n{}^N \leqslant \varphi(r_n{}^N). \tag{7}$$

It is known (see, e.g., [3]) that, for any sequence Λ,

$$\| s^{\lambda_N} - [\Lambda_N] \|^2 = \frac{1}{1+2\lambda_N} \prod_{n\neq N} \left(\frac{\lambda_n - \lambda_N}{\lambda_n + \lambda_N + 1}\right)^2, \tag{8}$$

where $\| \cdot \|$ denotes the L^2 norm on (0, 1). Thus, by (7),

$$\begin{aligned} \| s^{\lambda_N} - [\Lambda_N] \| &= (1+2\lambda_N)^{-1/2} \exp\left[-\sum_n \varphi_n{}^N\right] \\ &\geqslant (1+2\lambda_N)^{-1/2} \exp\left[-\sum_n \varphi(r_n{}^N)\right]. \end{aligned} \tag{8'}$$

LEMMA 1. *Let Λ, etc., be as above. Then*

$$\| s^{\lambda_N} - [\Lambda_N] \| \geqslant (1+2\lambda_N)^{-1/2} \exp[-N\varphi(r_N) - a_N\lambda_N\tau_N] > 0 \tag{9}$$

where

$$a_N = \max\{\varphi'(0), \varphi(r_N)/r_N\}.$$

PROOF. We break up the sum on the right of (8′) into $\sum_{n<N}$ and $\sum_{n>N}$. Then, as φ is an increasing function of r,

$$\sum_{n<N} \varphi(r_n{}^N) \leqslant N\varphi(r_N).$$

From the form of the function φ we have

$$\varphi(r) \leqslant a_N r \qquad (0 < r \leqslant r_N)$$

so that

$$\sum_{n>N} \varphi(r_n{}^N) \leqslant \sum_{n>N} a_N r_n{}^N = a_N \lambda_N \tau_N \,.$$

Combining these inequalities with (8′) gives (9).

LEMMA 2. *Let $P \in [\Lambda_N]$ so that $(s^{\lambda_N} - P) = \sum_n b_n s^{\lambda_n}$ with all but finitely many b_n zero and with $b_N = 1$. Then*

$$| b_n | \leqslant \| s^{\lambda_N} - P \| / \| s^{\lambda_n} - [\Lambda_n] \| \qquad (n = 0, 1, \ldots). \tag{10}$$

PROOF. Observe that (10) is immediate if $n = N$ or if $b_n = 0$. Otherwise $(n \neq N,\ b_n \neq 0)$, note that

$$\| s^{\lambda_N} - P \| = \left\| \sum_\nu b_\nu s^{\lambda_\nu} \right\| = | b_n | \left\| s^{\lambda_n} - \sum_{\nu \neq n} \frac{b_\nu}{b_n} s^{\lambda_\nu} \right\|$$

$$= | b_n | \, \| s^{\lambda_n} - P_0 \|$$

where

$$P_0 = \sum_{\nu \neq n} \frac{b_\nu}{b_n} s^{\lambda_\nu} \in [\Lambda_n].$$

Thus,

$$\| s^{\lambda_N} - P \| \geqslant | b_n | \, \| s^{\lambda_n} - [\Lambda_n] \|$$

which is just (10).

Note that by (6) the sequence Λ satisfies:

$$\sum_n \alpha^{\lambda_n} < \infty \qquad \text{for} \qquad 0 < \alpha < 1. \tag{11}$$

We now add the assumption that Λ satisfies the additonal condition:

$$\left[a_N \tau_N + \frac{N}{\lambda_N} \varphi(r_N) \right] \to 0 \qquad \text{as} \qquad N \to \infty. \tag{12}$$

LEMMA 3. *Let Λ, etc., be as above with Λ satisfying* (12). *Set*

$$\Gamma = \sum_0^\infty \alpha^{\lambda_n} \exp[n\varphi(r_n) + a_n \lambda_n \tau_n].$$

Then, for $0 < \alpha < 1$, the series Γ converges and, for $P \in [\Lambda_N]$,

$$\| s^{\lambda_N} - P \|_0 \leqslant \sqrt{\alpha}\, \Gamma \| s^{\lambda_N} - P \| \tag{13}$$

where $\| \cdot \|_0$ is the L^2 norm on $(0, \alpha)$.

PROOF. Let $\alpha < \alpha_1 < 1$. By (12), for large enough n we have

$$a_n \tau_n + \frac{n}{\lambda_n} \varphi(r_n) \leqslant \log \frac{\alpha_1}{\alpha}$$

so that

$$\alpha^{\lambda_n} \exp[n\varphi(r_n) + a_n \lambda_n \tau_n] = \exp\left[\lambda_n \left(\log \alpha + \alpha_n \tau_n + \frac{n}{\lambda_n} \varphi(r_n)\right)\right]$$
$$\leqslant \alpha_1^{\lambda_n}$$

and, by (11), the series converges. Observe, now, that

$$\| s^{\lambda_N} - P \|_0 = \left\| \sum_0^\infty b_n s^{\lambda_n} \right\|_0 \leqslant \sum_0^\infty | b_n | \, \| s^{\lambda_n} \|_0 .$$

Using (10) and then (9) and the evaluation

$$\| s^\lambda \|_0^2 = \alpha^{2\lambda+1}/(2\lambda + 1),$$

we obtain

$$\| s^{\lambda_N} - P \|_0 \leqslant \sum_0^\infty [\alpha^{2\lambda_n+1}/(2\lambda_n + 1)]^{1/2} \| s^{\lambda_N} - P \| / \| s^{\lambda_n} - [\Lambda_n] \|$$

$$\leqslant \sqrt{\alpha} \| s^{\lambda_N} - P \| \sum_0^\infty \alpha^{\lambda_n} \exp[n\varphi(r_n) + a_n \lambda_n \tau_n]$$

which is just (13).

We are now ready, at last, to obtain the desired lower bound for $| s^{\lambda_N} - [\Lambda_N] |$.

THEOREM 1. *Let $\Lambda = (\lambda_0, \lambda_1, \ldots)$ be an increasing sequence $(0 \leqslant \lambda_0 < \lambda_0 < \cdots)$ of reals satisfying (6) and (12) and let $0 < \alpha < 1$ be such that $(1 - \alpha \Gamma^2) = c^2 > 0$ (observe that $c = c(\alpha) \to 1$ as $\alpha \to 0$). Then*

$$| s^{\lambda_N} - [\Lambda_N] | \geqslant c \| s^{\lambda_N} - [\Lambda_N] \|$$
$$\geqslant c(2\lambda_N + 1)^{-1/2} \exp[-N\varphi(r_N) - a_N \lambda_N \tau_N] > 0. \tag{14}$$

PROOF. For any $P \in [\Lambda_N]$ we have

$$\| s^{\lambda_N} - P \|^2 = \| s^{\lambda_N} - P \|_0^2 + | s^{\lambda_N} - P |^2$$

by the definition of the norms. Using (13) gives

$$| s^{\lambda_N} - P |^2 \geqslant (1 - \alpha \Gamma^2) \| s^{\lambda_N} - P \|^2 \geqslant c^2 \| s^{\lambda_N} - [\Lambda_N] \|^2.$$

Since this holds for every $P \in [\Lambda_N]$,

$$| s^{\lambda_N} - [\Lambda_N] | \geqslant c \| s^{\lambda_N} - [\Lambda_N] \|$$

which, with (9), gives (14).

3. In order to apply the Theorem above to the prediction problem described earlier it is only necessary to show that the sequence $\Lambda = (n^2 : n = 0, 1,...)$ satisfies (6) and (12). Certainly (6) holds and, in fact,

$$\tau_N = \sum_{N+1}^{\infty} n^{-2} = \mathcal{O}(N^{-1}). \tag{15}$$

We have

$$\varphi(r) = | \log(1 - r)/(1 + 2r) | , \qquad r_n = n/(n + 1)$$

so, for large N, it follows that $(1 - r_N) = \mathcal{O}(N^{-1})$,

$$\varphi(r_N) = \mathcal{O}(\log(1 - r_N)) = \mathcal{O}(\log N) \tag{16}$$

and

$$a_N = \varphi(r_N)/r_N = \mathcal{O}(\log N). \tag{17}$$

Thus, combining (15), (16) and (17),

$$\begin{aligned} a_N \tau_N + (N/\lambda_N)\, \varphi(r_N) &= \mathcal{O}(N^{-1} \log N + (N/N^2) \log N) \\ &= \mathcal{O}(N^{-1} \log N) \end{aligned}$$

which goes to 0 as $N \to \infty$, satisfying (12).

It follows, therefore, that Theorem 1 may be applied. Hence for α sufficiently small so that

$$\alpha \Gamma^2 < 1, \tag{18}$$

we have $c^2 = 1 - \alpha\Gamma^2 > 0$ and thus, by (5) and (14),

$$\begin{aligned} \| l_N \| &\leqslant c^{-1}(2N^2 + 1)^{1/2} \exp[N\varphi(r_N) + a_N \lambda_N \tau_N] \\ &= \exp[\mathcal{O}(N \log N)]. \end{aligned} \tag{19}$$

Substituting this in (3) we see that the factor $\alpha^{2N^2} = \exp[-\mathcal{O}(N^2)]$ dominates and the series converges. We have thus shown the following.

THEOREM 2. *For T large enough that* (18) *is satisfied with $\alpha = e^{-\pi^2 T}$, the mapping*

$$A : f = u(\cdot, 0) \mapsto u(T, \cdot) = w$$

is a well-defined, bounded (using L^2 norms) linear map for solutions u of (1) *with $0 < t < T$; i.e., the "observation and prediction" problem is well-posed.*

REMARK. The above proof does not, of course, show that A is necessarily unbounded for smaller $T > 0$. It would seem of interest to determine whether this notion of a "minimal period of observation" is, indeed, a genuine phenomenon or whether it is, perhaps, imposed merely by the exigencies of this particular method of proof. Since $\Gamma \to \infty$ as $\alpha \to 1$, (18) is genuinely a restriction on T and it would also be of interest to estimate the minimal T for which (18) is satisfied. The complications involved in estimating the sum for Γ make the task of estimating this number forbiddingly difficult.

4. A couple of generalizations of Theorem 2 suggest themselves. For example, it is clear that the similar problem for a non-uniform bar

$$u_t = a(x)\, u_{xx}, \qquad u_x(t, 0) = u_x(t, 1) = 0 \tag{1'}$$

could be treated the same way using the expansion

$$u(t, x) = \sum_n c_n e^{-\lambda_n t} v_n(x) \tag{2'}$$

where the $\{\lambda_n\}$ and $\{v_n\}$ are the appropriate eigenvalues and eigenfunctions. The asymptotic behavior of the $\{\lambda_n\}$ is known to be similar enough to that of $\{n^2\}$ to ensure the applicability of Theorem 1. The only new feature introduced would be the estimation of an asymptotic lower bound for $\{v_n(0)\}$.

Another example is the treatment of the observation and prediction problem for a solid body, rather than a rod. For a cylindrical body in R^k we have the following result.

THEOREM 3. *Let $\mathscr{D}$ be a "suitable" region in R^{k-1} and let $\mathscr{D}_* = (0, 1) \times \mathscr{D}$.*

Then, for T large enough (as in Theorem 2), the mapping from $L^2((0, T) \times \mathscr{D})$ to $L^2(\mathscr{D}_*)$ defined by

$$A : f = f(t, y) = u(t, 0, y) \mapsto w = w(x, y) = u(T, x, y)$$

is a well defined, bounded (using L^2 norms) linear map for solutions u of

$$u_t = \Delta u \; (= u_{xx} + \Delta_y u) \qquad (0 < t < T, (x, y) \in \mathscr{D}_*),$$

$$\frac{\partial u}{\partial \nu} = 0 \qquad (0 < t < T, (x, y) \in \partial\mathscr{D}_*). \tag{1''}$$

PROOF. We use the expansion

$$u(t, x, y) = \sum_{m,n} c_{m,n} \exp[-(n^2 + \mu_m)\, \pi^2 t]\, v_m(y) \cos n\pi x \tag{2''}$$

where the $\{v_m\}$ are normalized solutions of

$$\Delta_y v_m + \mu_m \pi^2 v_m = 0 \qquad (y \in \mathscr{D}),$$

$$\frac{\partial v_m}{\partial \nu} = 0 \qquad (y \in \partial\mathscr{D}).$$

Since the $\{v_m\}$ form an orthonormal sequence we have

$$\langle u(t, 0, \cdot), v_m \rangle_{\mathscr{D}} = e^{-\mu_m \pi^2 t} \sum_n c_{m,n} e^{-n^2 \pi^2 t}.$$

Hence, letting l_N be as originally, we have

$$c_{m,N} = l_N(s^{-\mu_m} \langle \varphi(s, \cdot), v_m \rangle_{\mathscr{D}}),$$

where

$$\varphi(s, y) = f(t, y) = u(t, 0, y)$$

on making the substitution $s = e^{-\pi^2 t}$ as before. Thus

$$Af = \sum_{m,n} l_n(s^{-\mu_m} \langle \varphi(s, \cdot), v_m \rangle_{\mathscr{D}})\, \alpha^{n^2 + \mu_m} v_m(y) \cos n\pi x$$

and

$$\| Af \|^2 \leqslant \sum_n \alpha^{2n^2} \| l_n \|^2 \sum_m \alpha^{2\mu_m} \,|\, s^{-\mu_m} \langle \varphi, v_m \rangle_{\mathscr{D}} \,|^2$$

$$\leqslant \sum_n \alpha^{2n^2} \| l_n \|^2 \sum_m |\, \langle \varphi, v_m \rangle_{\mathscr{D}} \,|^2$$

$$= \sum_n \alpha^{2n^2} \| l_n \|^2 \| \varphi \|^2_{(\alpha,1) \times \mathscr{D}} \leqslant \pi^2 \sum_n \alpha^{2n} \; \| l_n \|^2 \| f \|^2_{(0,T) \times \mathscr{D}} \,.$$

This gives the identical estimate (3) for $\| A \|$ as in Theorem 1 and the same proof now goes through.

The result of Theorem 3 suggests a conjecture that a similar prediction problem would be well-posed given observation, for a sufficiently long period, of the restriction to a nonempty relatively open subset $\Omega \subseteq \partial\mathscr{D}_*$ where $\mathscr{D}_*$ is a more general domain in R^n (Ω here corresponds to $\{0\} \times \mathscr{D}$ in Theorem 3). The methods of this paper, however, seem to afford no direct mode of attack on this more general problem; the proof of Theorem 3 makes essential use of the special nature of $\mathscr{D}_*$ and Ω.

Note added in proof: Use of Theorem II of § 9 of [4] immediately shows that the inequality (14) holds for all $\alpha > 0$, and hence that the conclusions of Theorems 2 and 3 hold for all $T > 0$ (i.e., there is no minimum observation time).

Acknowledgment

This research was supported by the National Science Foundation under Grant GP 7607 and by the Air Force Office of Scientific Research under Grant AFOSR 62-414.

References

1. J. A. Clarkson and P. Erdös. Approximation by polynomials. *Duke Math. J.* **10** (1943), 5–11.
2. R. C. MacCamy, V. J. Mizel, and T. I. Seidman. Realization of temperature distributions. *J. Math. Anal. Appl.* (1968).
3. W. Rudin. Real and complex analysis. McGraw-Hill, New York, 1966.
4. L. Schwartz, "Étude des sommes d'exponentielles," 2nd ed. Hermann, Paris, 1959.